Computing Summary

F. Bianchi
Torino

V SuperB Collaboration Meeting
Pisa, September 21st, 2012

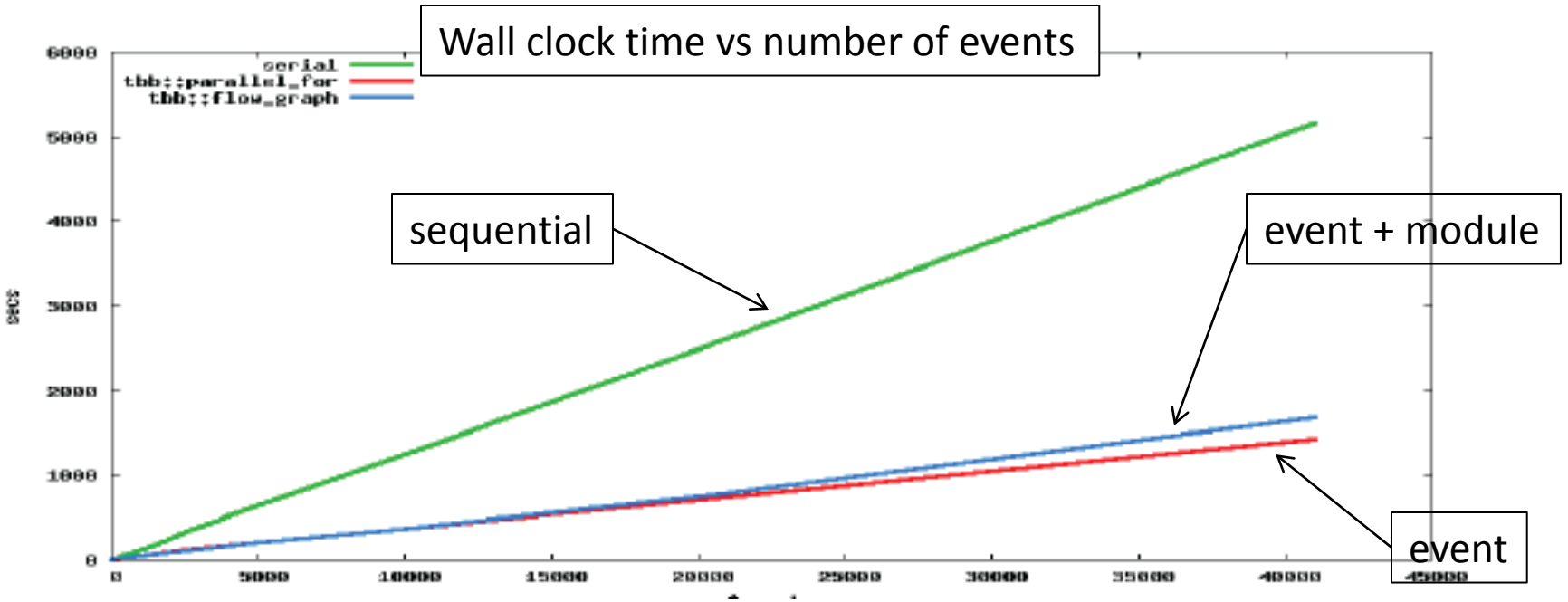
UNIVERSITÀ
DEGLI STUDI
DI TORINO
ALMA UNIVERSITAS
TAURINENSIS

Outline

- Lively parallel sessions!
  - R&D
  - Distributed Computing Tools
  - FullSim & Background
  - Computing section of Detector TDR
- My selection of presented material plus some comments
  - Look at the presentations in Indico for details

R&D: Parallel Framework (1)

- M. Corvo has presented nice progresses.
- Test were made using the current FastSim Framework that is based on the BaBar code.
- Test both of event level and event plus module parallelism.



R&D: Parallel Framework (2)

- We need to come to a full understanding of all dependencies among modules
- This is a good starting point, but there are a lot of analysis patterns to consider, not to mention a full reconstruction framework
- This exercise was helpful to understand also what modules can do and what modules must NOT do to improve optimization
- There's room for improvements also in the generation part, trying to limit (or at least factorize) the use of Common Blocks in Fortran code
- In the long term we will abandon the current framework for a new one which is natively parallel and whose architecture will be designed according to our experiences

R&D: Distributed Storage

- Storage system evaluation
  - HadoopFS on WAN: testbed on Bari and Napoli
- Data access framework library development
  - Data access optimization on local and WAN scenario
  - Mask the low level data access layer at the sites
  - Useful support from ROOT development team
- Mass data transfer system
  - FTS3 evaluation
  - PhEDEx evaluation process (standby)
- File catalog ng (dynamic LFC ng by EMI R&D, standby)

DIRAC (1)

- DIRAC (*Distributed Infrastructure with Remote Agent Control)* is a Grid resource management tool.
  - Goal: setup and configure a Dirac system to fulfil the SuperB requirements
  - General work plan:
    - Simulation production use case (in progress)
    - Workload Monitor system
    - Analysis use case integrated with Ganga system
    - Mass data transfer system

DIRAC (2)

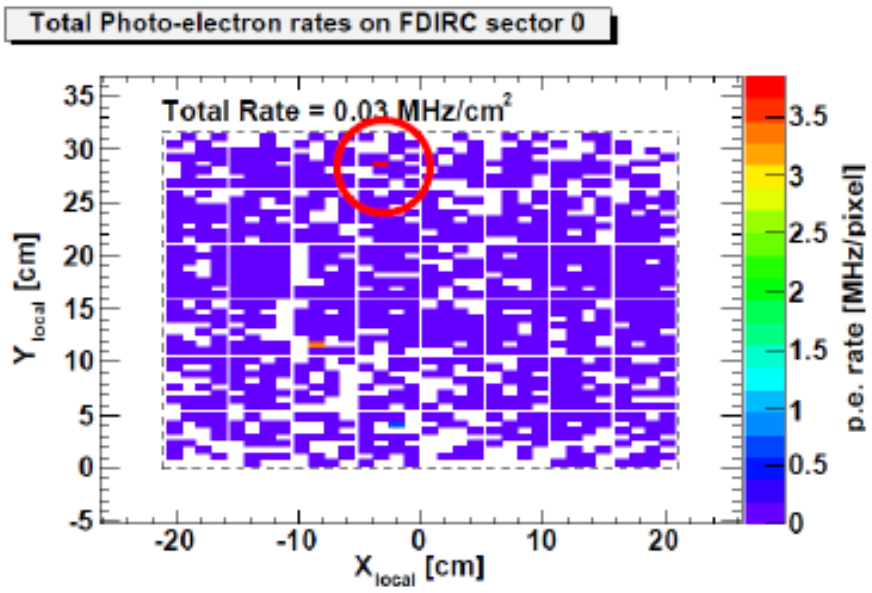
- Present goal: manage all aspects related to Simulation Production (FastSim and FullSim) via DIRAC
  - User management (role and permissions)
  - Site Management per Session
  - Production creation and monitoring
  - Requests creation and monitoring
  - Bunch jobs submission and monitoring
  - DIRAC need to interact with Bookeeping Database (SBK5)
- Test to launch Severus (python wrapper for SuperB simulation production) via DIRAC have been successfully performed
- Very productive interactions with the DIRAC development team

FullSim & Background

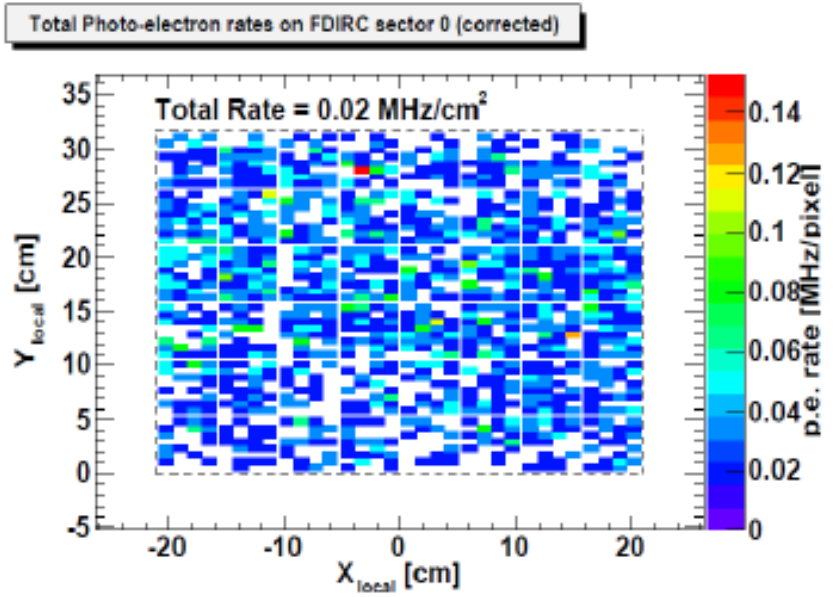
- Summer production was a success
  - A very complete set of background samples have been produced
    - Rad-bhabha (low and high k)
    - Pairs
    - Touschek and BeamGas (HER/LER)
    - Synchrotron Radiation (HER/LER)
  - Several improvement in detector and B field model
- Production was affected by a bug on optical photons in FDIRC
  - Luckily was recoverable offline
  - Need to improve our QA protocol before starting production
  - More advance planning of production is needed
- Good stress test of new distributed production system
  - Transfer of large (>3 GeV) files over the Grid failed.

The patch in action

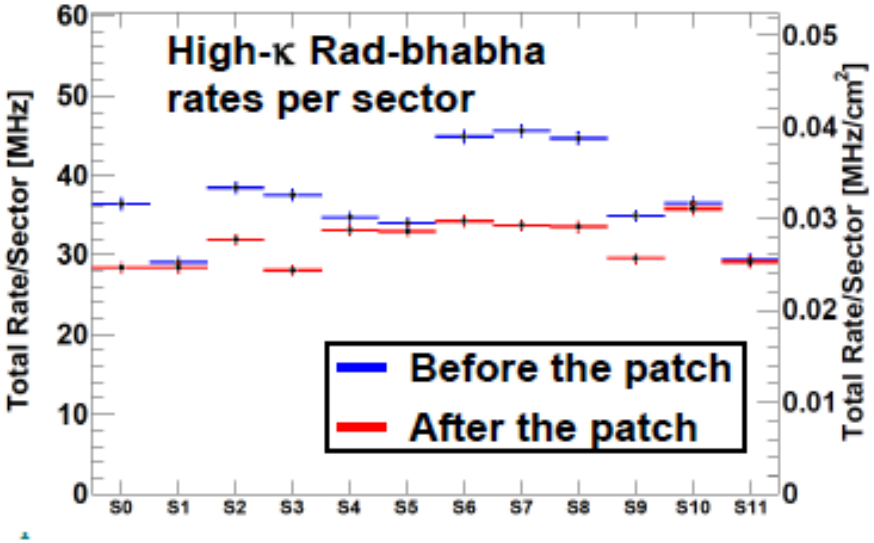
Before the patch



After the patch



- Algorithm correctly identify “hot pixels” and filter them out
- The total rates in some sectors get significantly reduced by up to 25-30%



Next productions

- Currently we are exploiting resources on the Grid for central production(s).
  - But output is sent back to CNAF and analyzed there.
- We should start to use resources on the Grid also for analysis.
- We have a Ganga based tool to submit analysis job on the Grid.
  - It has been tested by few volunteers
  - Should become of general use sometime early next year.
  - Our plan is to offer a tutorial at the first 2013 meeting and ask people to use it for their analysis soon after.

TDR Writing Recommendation

- Please remember that the repository is case sensitive.
  - The files "myfile.pdf" and "myfile.PDF" are different files.
- Please avoid committing to the repository a "dtdr.tex" file with all subsystems (with the exclusion of yours) commented out.
- Please resolve all conflicts before committing the text to the repository
  - Do a "svn update"
  - Resolve all conflicts by editing the affected files
  - Finally do the "svn commit"
- Please avoid lines with >1024 characters
- Check with the "svn stat" command
  - Files marked with a "c" have conflicts
  - Files marked with a "?" are missing

Computing Section of DTDR

Status

- All material is there
  - Only final summary is missing
- More work by the computing group to improve wording is needed.
  - Plan is to complete this step in the next couple of weeks

Conclusions

- The computing group is supporting the collaboration by providing:
  - Collaborative Tools
  - Physics Tools: FastSim, etc.
  - FullSim
  - Production Tools
  - Bookkeeping Tools
- There is an active R&D program aimed at the design of the computing model.
- The activities funded under the Pon ReCaS are an important step forward into building the computing infrastructure.
- A severe lack of manpower is affecting us.
- Come and join the fun !